

BorgBackup

„The holy grail of backup software“

Thomas Waldmann @ 32c3 (Dec 2015)

BorgBackup = a fork of Attic Backup

- attic: 5y old, good design, proven code
 - but: slow-going / stalled development
 - not very open to new developers
- borg: forked ~ 5 / 2015
 - fast-paced development
 - open and inviting to developers

Features (1)

- easy, single-file binary, no installation
- compression: lz4, zlib, lzma
- authenticated encryption (ETM AES256 + HMAC)
- backups to:
 - mounted filesystem (local or network)
 - ssh server with borg installed
- content-defined chunking
- chunk deduplication

Features (2)

- FOSS (BSD license), good docs
- platforms: Linux, *BSD, OS X, [cygwin]
- archs: tested on x86, x86_64, arm v7
- xattr / acl support, hardlinks, ...
- mount backups via FUSE
- good test coverage, CI
- Python 3.x + a little Cython and C

Deduplication (1)

- No problem with:
 - VM images (simple sparse file support)
 - (physical) disk images
 - renames of huge directories
 - inner deduplication of data set
 - historical deduplication
 - dedups files on different machines

Deduplication (2)

- Content Defined Chunking
 - „buzhash“ rolling hash
 - $h(\text{data}[\text{pos} : \text{pos} + \text{window size}]) \& \text{MASK} == 0 \rightarrow \text{cut!}$
 - yields pieces of $\sim 2^{\text{MASK}}$ bytes size
 - chunker params configurable
 - seeded to avoid fingerprinting via chunk lengths
- Stores chunks into kv store:
 - $\text{key} = \text{HASH}(\text{chunk})$ or: $\text{HMAC}(\text{key}, \text{chunk})$
 - $\text{value} = (\text{compressed} / \text{encrypted} / \text{auth.}) \text{ data}$

Links

github.com/borgbackup

Questions / Feedback?

- meet me at the Python assembly
- @ThomasJWaldmann on Twitter
- #borgbackup on chat.freenode.net
- use Github :)